DTU Library

Updating distributed hydrodynamic urban drainage models

Borup, Morten; Grum, Morten; Madsen, Henrik; Mikkelsen, Peter Steen

Published in:
Proceedings of the 13th International Conference on Urban Drainage

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Updating distributed hydrodynamic urban drainage models

Morten Borup[1*], Morten Grum[2], Henrik Madsen[3] and Peter Steen Mikkelsen[1]

[1] *Department of environmental Engineering. Technical University of Denmark*
[2] *Krüger A/S, Veolia Water Solutions and Technologies, Denmark*
[3]*DHI, Denmark*
**Corresponding author's e-mail: morb@env.dtu.dk*

ABSTRACT
Online distributed urban drainage models (DUDMs) can be useful for aiding real time control decisions as well as for issuing flood warnings and assessing whether or not overflow occurs from ungauged CSOs. For an online model to be useful it is required that it can be updated to current conditions by data assimilation. In the current work two existing deterministic methods for updating DUDMs are described as well as a novel implementation of the Ensemble Kalman Filter (EnKF) for these models. While the deterministic methods are the most computationally efficient, they cannot explicitly update the entire model, cannot produce uncertainty estimates and cannot treat ambiguous information from multiple sensors in a meaningful way. All this is possible with the EnKF but at the expense of a significant computational overhead. In a synthetic and a real world experiment the EnKF was shown to be able to update the entire models based on as well upstream water level measurements as downstream flow observations. The conclusion is that the EnKF is to be preferred over the deterministic methods if the required computational resources are available.

KEYWORDS
Distributed urban drainage models; data assimilation; Ensemble Kalman Filter; online models.

INTRODUCTION
The modern distributed hydrodynamic urban drainage models (DUDMs) such as MIKE URBAN, SWMM, InfoWorks etc., contain an overwhelming amount of information from the drainage system. The hydrodynamic part of these models is based on physical data, such as pipe location, depth, length and diameter as well as detailed descriptions of the actuators in the system, such as gates, weirs and pumps. This enables the models to predict the behaviour of the system without dependency on prior training or calibration, which can be a major advantage for real time control purposes for ever changing systems like the modern drainage system. The potential uses of online DUDMs are not limited to real time control. Other examples could be predicting local flooding, surveillance of system (mis-)behaviour due to e.g. sedimentation, and model based cross validation of gauges which can contribute to a higher level of data security and save maintenance costs.

For an online model to be useful it is necessary to be able to adapt the model to current conditions indicated by system measurements. This process is in the following referred to as *data assimilation* (DA) or *updating*. Data assimilation has for long been widely used within natural catchment hydrology but this has not been the case within distributed urban drainage modelling. The reasons for this are manifold. First of all the need for knowing the state of the system in real time might not have been obvious in the past since the drainage systems were less controlled and the level of ambition in terms of system performance probably lower than today. Secondly, the technology in terms of computational power and online data acquisition has been inadequate. Today both the need and the technology are there. The current work elaborates on the options, challenges, requirements and methods for updating DUDMs. Two existing previously documented deterministic DA methods are briefly described. These are deterministic in the sense that they update model variables to a single deterministic value without explicitly considering uncertainty. These methods were developed solely for DUDMs as a way of circumventing the rather large computational overhead required for the more formal statistically based methods, such as e.g. the Ensemble Kalman Filter (EnKF) (Evensen, 2003, 1994). The statistically based methods do have some potential large benefits, however, and therefore the main DA method used in this article is the EnKF. The EnKF was first developed for oceanography and has since been used with success within various research areas such as geo-physics (Aanonsen et al., 2009) and hydrological stream flow forecasting (Clark et al., 2008; Rakovec et al., 2012) but to the authors' knowledge this is the first time in the open literature this has been used to update DUDMs. Simple synthetic setups as well as an actual full scale DUDMs are used for illustrations.

DETERMINISTIC UPDATING FOR DUDMS

Point wise updating

The simplest way of updating a model is to force individual observed model variables to equal measurements. This means that in the case of continuous measurements this work as a kind of temporary boundary condition, that define a specific model variable until the last observation. The updating only has a direct effect on observed variables but the normal hydrodynamic computations make sure that the impact of the corrections propagates to at least the more downstream parts of the model. Therefore, for this updating method to have a system wide effect, it is necessary to include a lot of observed locations in the updating and many of these locations should be as far upstream as possible. Since the method does not estimate the uncertainty of model variables it cannot perform a dynamic weighting between model and observation uncertainty in the updating. Therefore well determined model estimates can end up being replaced by uncertain observations, or vice versa.

The commercially available point wise deterministic updating that is part of the Mike Urban software suite was tested by Hansen *et al.* (2011). In this study the water levels in 8 different basins in a DUDM of a 90.000 inhabitant Danish city were updated. This improved model forecasts substantially for the receding hydrograph after an event.

Inflow correction

There is a large degree of correlation in most hydrological systems; if the downstream flow rate is higher than normal it is very likely that upstream flow rates are higher than normal as well. Furthermore, due to the conservation of mass in the hydrodynamic models, the flow into and out of any part of the model must be the same over time. Based on these facts Borup *et al.* (2011) presented a method for correcting the states of the surface models that governs the inflow to the hydrodynamic part of DUDMs based on the difference between model and observed downstream flow. The core of this method is to use a linearised version of the model to keep track of previous corrections in order to enable for fast corrections to the model states without provoking model instabilities. Besides the initial estimation of the linearised version of the hydrodynamic model the method does not require any computational overhead. The method is, however, limited to correcting the states of the surface module that governs processes of time scales larger than the hydraulic response time of the network model, which means that it foremost is suited for updating infiltrating inflow and base flow.

Everything upstream of the flow gauge is updated in a lumped manner and the network model is only updated implicitly as the updated inflow propagates down through the model. Therefore the method is not suitable for including multiple gauges in the same catchment.

ENSEMBLE BASED UPDATING

The classical Kalman filter (Kalman, 1960) updates model states from the difference between model and measurements, which in the following is called the innovation, as:

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \boldsymbol{K}\boldsymbol{i}$$

where $\boldsymbol{x}^a$ and $\boldsymbol{x}^b$ is the state vector after and before an update, respectively, $\boldsymbol{K}$ is the Kalman gain matrix that determines how much each state variable should be updated from given innovations and $\boldsymbol{i}$ is a vector containing the innovations for all observations. In the classical Kalman filter $\boldsymbol{K}$ is calculated from the full error covariance matrix that is propagated forward in time using the linear model operator, or a linearised model operator in the case of a non-linear model.

The Ensemble Kalman Filter is a Monte Carlo implementation of the classical Kalman filter in which an ensemble of models is propagated in parallel to estimate the model error statistics. Even though this sounds like a large computational overhead it is actually a dramatic reduction in computational cost compared to the classical Kalman filter, when system dimensions are large. The Kalman filter requires the computation of the covariance between all state variables in the model, which result in an n by n matrix where n is the system dimension, and for DUDMs the system dimensions easily supersede 10^4. Therefore the parallel propagation of an ensemble of 10 to 100 models is comparably very cheap. Besides being efficient for large models the EnKF is useful for nonlinear models due to the fact that it is the model itself that propagates the uncertainty statistics forward in time. More details on the EnKF can be found in (Evensen, 2003).

State vector

One has to define the DA state vector to be updated by the EnKF. The St. Venant equations that are the basis of the hydrodynamic computations of DUDMs include both discharge and water level variables. Changes to the discharge variables have no lasting influence on the model, however, since they do not determine the amount of water in the system. Therefore it was chosen only to update the model variables that alter the amount of water in the model: i.e. the water levels in the hydrodynamic model and the storages in the linear reservoirs of the surface module.

Filter noise

The ensemble of models that is the backbone of EnKF is formed and maintained by affecting it with stochastic realizations of all major uncertainties affecting the updated model. For DUDMs the main source of uncertainty is the rainfall data used to force the model. For distributed models the uncertainty description should be in both space and time, which means that it is preferable to use rainfall data which includes a spatial description of the rainfall, i.e. radar data. The quantitative quality of rainfall estimates from radar might often be lower than those from rain gauges, but if the relative distribution of the rainfall is described correctly it will enable for efficient updating that can make up for errors caused by the rainfall estimates. Besides, it was shown by Borup *et al.* (2013) that the negative impact of the time displacements induced in models by using raingauge data, due to the time it takes for rain cells to travel the distance between gauge and catchments, can justify the use of radar data of much lower quantitative quality. For these reasons it has been decided to use a noise formulation for the rainfall data ensemble that is inspired by the authors' perception of the statistical properties of radar rainfall estimation errors. In short this includes multiplying the rainfall estimate for each ensemble member with a factor f that is drawn from a uniform distribution *unif*(0.1. 1.9) every *t* minutes, where *t* itself is a random variable that is drawn from *unif*(1, 60) every time a new f is drawn. The rainfall estimates are assumed to be uniformly distributed in all examples for the sake of simplicity.

In order to reflect the uncertainty of the inflow to the pipe system when it is not raining, the following formulation is used to perturb the linear reservoirs of the surface models:

$$x'_k = x_k + \sqrt{1-\beta^2}\, w_{k-1} em_k, \quad \beta = e^{-\Delta t/\tau}$$

where *k* is the time step index, *x* is the individual surface model state variable, *w* is zero mean white noise with variance σ_w^2, *em* is the ensemble mean and τ is the time constant of the linear reservoir model.

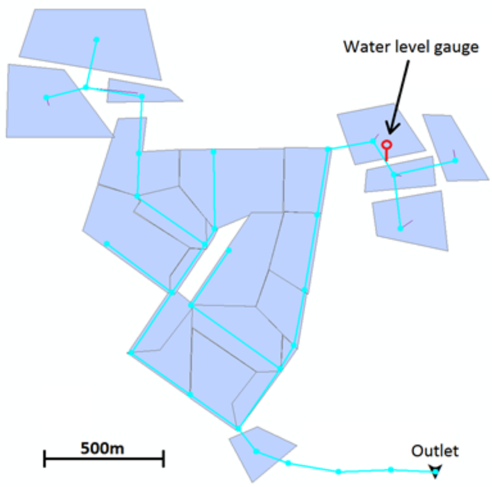
Choice of EnKF flavour and ensemble size

In the standard EnKF formulation also the uncertainty of the system observations have to be represented by stochastic realisations, which result in additional sampling errors and spurious long range correlations for small ensemble sizes. In order to reduce these problems the formulation of the EnKF by Sakov and Oke (2008) that do not rely on perturbed observations is used in the following. In the current work only ensemble sizes of 10 members are used.

CASE STUDIES

Utilizing upstream water level measurements with the EnKF

In this case study an upstream water level gauge is used to update an entire DUDM of a 107 ha catchment in a synthetic experiment using the EnKF. In the experiment an 11 day rainfall series is used as the true rainfall which is propagated through the model to create the true model states. Rainfall and water level observations are made by perturbing the true rainfall and true model states, respectively. The quality of the updated model is assessed by comparing the modelled flow out of the catchment with the true flow. A view of the model can be seen along with the definition of rainfall perturbations used for the test in Figure 1. The different rainfall scenarios are used to illustrate the impact of the quality of the rainfall estimate on the updated model.



Rainfall scenario 1

Rain data perturbed as described in the "Filter noise" section.

Rainfall scenario 2

Estimated rainfall intensity of 2.5 um/s during rain, otherwise 0 um/s.

Figure 1. View of the distributed model. The pipe system is highlighted in magenta.

Gaussian white noise observation errors with a standard deviation of 10 cm has been used for creating the water level observations as well as for the filter configuration. To test the impact of unacknowledged observation bias the test was also performed using observations that were affected with a 5 cm positive bias. The results of the tests can be seen in Figure 2. The results clearly shows that the EnKF is capable of improving the quality of the model estimates by including data from a single water level gauge from a remote corner of the system. The results also show that even the error caused by very pour rainfall estimates can be corrected by a single upstream water level gauge, when the rainfall estimation error is spatially homogeneous as in the current example.

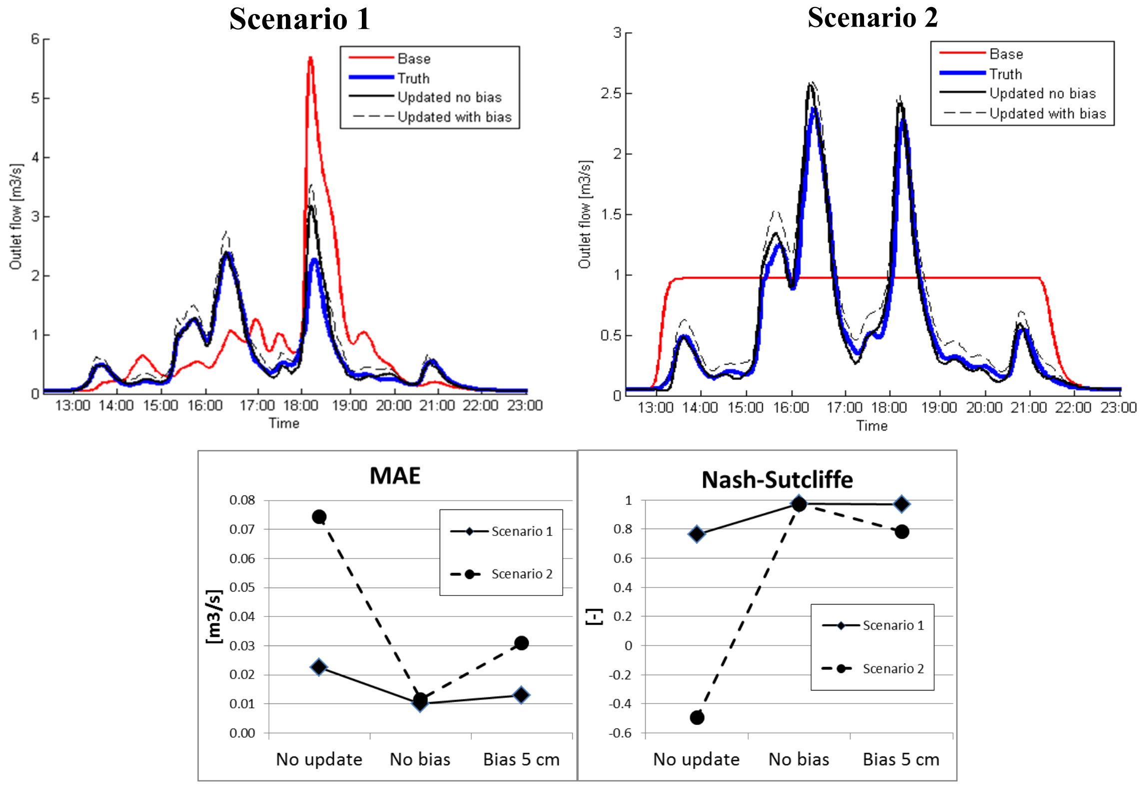


Figure 2. (Top panel) Example of outflow hydrographs for an event for Scenario 1 (left) and Scenario 2 (right). The solid black line in the model results when the observation noise is zero mean (and std = 10 cm). The dotted black line occurs when observation noise has a bias of 5 cm, which is not accounted for by the filter. The red line is simulation without updating. (Lower panel) Mean Absolute Error (MAE) and Nash-Sutcliffe efficiency index (Nash and Sutcliffe, 1970) for the entire 11-day period.

Real case study

In the following the EnKF is used to update an actual full scale DUDM of a 2000 ha urban catchment based on downstream flow measurements. Since the EnKF uses the covariance between model states and the innovations for updating, it is not required that the observed quantities are of the same type as the measured quantities. This allows for the current setup where discharge observations are used to update water levels and not discharge. The same catchment (see Figure 3) and data were used by Borup *et al.*(2011) to test deterministic inflow correction of slow changing flow components, since there is a lot of infiltrating water into the pipe system. During the validation period there was modelled overflow from several overflow structures, see Figure 3, which makes the model response very non-linear. The rainfall data comes from a single raingauge.

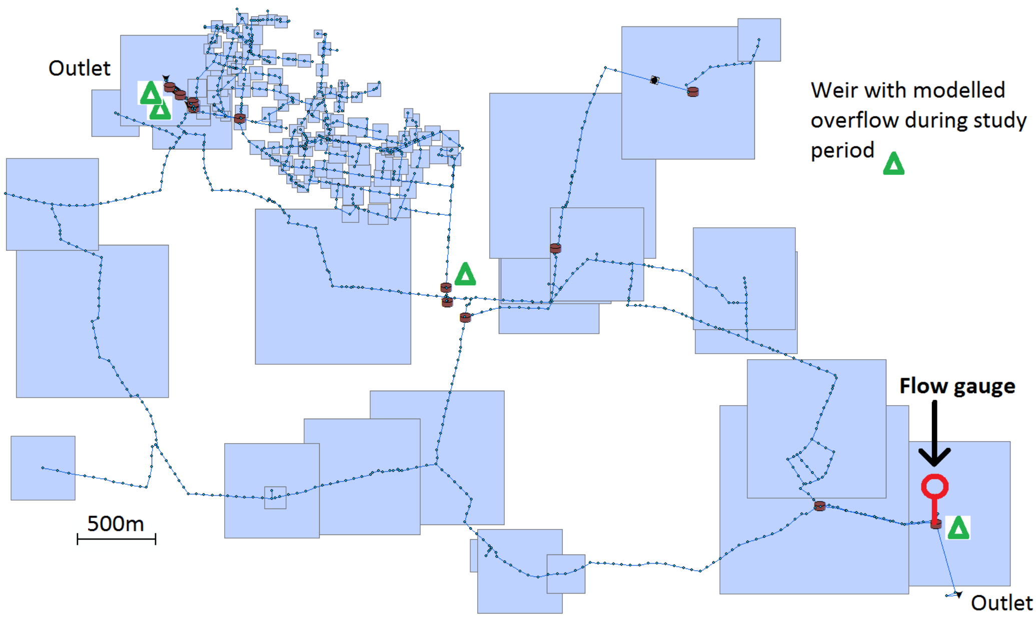


Figure 3. View of the Mike Urban model of the Ballerup catchment. The squares show the individual sub-catchments of the surface module.

The quality of the updates is assessed by creating forecasts that are initiated from the updated model and propagated up to ten hours ahead using the measured rainfall as model forcing. A correct update of very upstream parts of the model must then be expected to produce better long term downstream flow forecasts, while a correct update of the parts of the model that is close to the flow gauge will result in better short term forecasts. This has been quantified by the Nash-Sutcliffe R^2 in Figure 4 (right). On the same Figure the results are shown from Borup *et al.*(2011) were deterministic inflow correction is used on the same data. It is seen that the two methods converge for the long horizons, meaning that they are equally good at updating the slow changing inflow parts of the surface module, while the EnKF clearly performs better for the short horizons, which is easily explained by the fact that only this method updates the states of the hydraulic system.

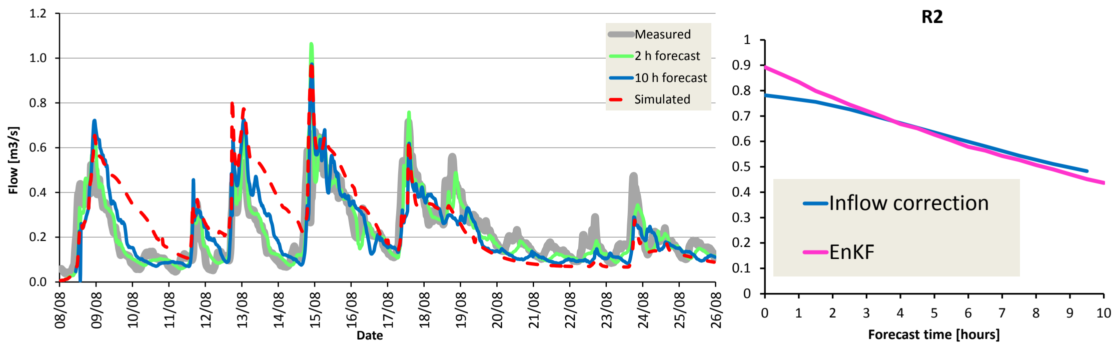


Figure 4. (Left) Observed discharge, simulated discharge without updating and the two and ten hour forecasts based on the DEnKF updated model. (Right) Nash-Sutcliffe efficiency index for forecasts initiated from a model that has been updated using the EnKF and deterministic inflow correction, respectively.

DISCUSSION AND CONCLUSIONS

Distributed urban drainage models (DUDMs) can be updated to reflect current conditions using various data assimilation (DA) methods. The choice of method depends upon the level of ambition as well as the properties of the system and model. The simplest DA method is point wise deterministic updating that forces individual variables in the model to equal the measurements. For this to have a notable effect on the model the states that are updated must have a big hydraulic impact on the model, which means that this method foremost is relevant for updating water levels in basins. In catchments were a major part of the runoff is routed through basins in wet weather this simple DA method can improve the downstream model estimates considerably without any computational overhead. Another deterministic method uses downstream flow measurements to correct the inflow to the hydrodynamic part of DUDMs and thereby only updating the hydraulic conditions in the network model implicitly. This method is efficient only for systems with a lot of slowly changing inflow, such as infiltrating water.

A much more general method that is capable of updating the entire model for any type of system is the Ensemble Kalman Filter (EnKF). The method is a Monte Carlo implementation of the standard Kalman filter and requires the parallel propagation of an ensemble of models, each of which are affected with perturbed model forcing and system noise to produce the correct model error statistics. This method can utilise both water level and discharge observations and is capable of using any number of observations in parallel. The method is not dependent upon training on historical data and can easily adapt to changes in the system behaviour due to change of actuator settings (pumps, gates, moveable weirs etc.) as long as these are included in the model. The only apparent drawback of the method is the computational requirements for propagating the ensemble of models forward in time. In the current work ensembles of 10 were used, which means that there is a computational overhead of a factor of 10 compared to just running a normal simulation. For this overhead one achieves the before mentioned benefits as well as uncertainty estimates on the model states. This means that the ensemble of models can be used to combine all observations from the system and give a quantitative assessment of the risk of flooding or overflow anywhere in the system.

In the current study the EnKF was used to update all mass balance related variables in DUDMs. It was shown in a synthetic experiment that measured water levels from a remote upstream corner of a catchment can be used to improve modelled discharge from the entire catchment greatly when the rainfall estimation error is spatially homogeneously distributed. In a real world experiment it was demonstrated that downstream discharge observations also can be used for updating DUDMs with the EnKF, and that the update improved model performance even though the rainfall was assumed to be spatially homogeneous, which is not a reasonable assumption for a model of this size. This shows that the EnKF can be used to update DUDMs even when the spatial distribution of the rainfall is not known.

The advantage of the deterministic data assimilation methods is that they imply little or no computational overhead. The disadvantages are that thet do not produce uncertainty estimates, that they cannot update the entire model explicitly, and that they cannot handle ambiguous information from multiple gauges in a meaningful way. The advantage of the EnKF data assimilation methods is that it produces uncertainty estimates, that it can update the entire model explicitly, and that it can handle ambiguous information from multiple gauges in a meaningful way. The disadvantage is the large computational overhead (>x10).

The overall conclusion is clear: If the required computational power is available the EnKF is the better way to update distributed urban drainage models.

ACKNOWLEDGEMENTS

This research was supported financially by the Danish Council for Strategic Research, Programme Commission on Sustainable Energy and Environment, as part of the Storm and Wastewater Informatics (SWI) project (http://www.swi.env.dtu.dk). Thanks to DHI for giving access to their data assimilation framework prototype and the API for the new Mike1D computational engine of Mike Urban.

REFERENCES

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Vallès, B., others, 2009. The Ensemble Kalman Filter in Reservoir Engineering--a Review. Spe Journal 14, 393–412.

Borup, M., Grum, M., Mikkelsen, P.S., 2011. Real time adjustment of slow changing flow components in distributed urban runoff models, in: Proceedings of the 12th International Conference on Urban Drainage.

Borup, M., Grum, M., Mikkelsen, P.S., 2013. Comparing the impact of time displaced and biased precipitation estimates for online updated urban runoff models. Water Science and Technology, 68(1), 109–16.

Clark, M.P., Rupp, D.E., Woods, R. a, Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., Uddstrom, M.J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Advances in Water Resources 31, 1309–1324.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research 99, 10143–10162.

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367.

Hansen, L.S., Borup, M., Møller, A., Mikkelsen, P.S., 2011. Flow Forecasting in Urban Drainage Systems using Deterministic Updating of Water Levels in Distributed Hydraulic Models, in: Proceedings of the 12th International Conference on Urban Drainage.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of basic Engineering 82, 35–45.

Nash, Je., Sutcliffe, J. V, 1970. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology 10, 282–290.

Rakovec, O., Weerts, a. H., Hazenberg, P., Torfs, P.J.J.F., Uijlenhoet, R., 2012. State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy. Hydrology and Earth System Sciences 16, 3435–3449.

Sakov, P., Oke, P.R., 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus A 60, 361–371.